

Limiting distribution of last passage percolation models

Jinho Baik*

February 1, 2008

Abstract

We survey some results and applications of last percolation models of which the limiting distribution can be evaluated.

1 Introduction

Consider a Poisson process of rate 1 in \mathbb{R}_+^2 . If one is only interested in the points in a square $(0, t) \times (0, t)$, it is equivalent to think of picking \mathcal{N} points at random in the square where \mathcal{N} itself is a random variable such that

$$\mathbb{P}(\mathcal{N} = N) = \frac{e^{-t^2}(t^2)^N}{N!}, \quad N = 0, 1, 2, \dots \quad (1)$$

Given a realization of the process, an up/right path from $(0, 0)$ to (t, t) is a piecewise linear curve starting from $(0, 0)$ to (t, t) joining Poisson points such that the slope, where defined, is positive. The length of an up/right path is defined by the number of points on it. Let $L(t)$ be the *length of the longest up/right path* from $(0, 0)$ to (t, t) , making it a random variable. See Fig. 1 for an example of a last longest up/right path. This is a type of last passage percolation model which is to find a path that maximizes the total weight (passage time) in a random environment.

Various statistics of $L(t)$ as $t \rightarrow \infty$ have been of interest. The basic law is[4]

$$\lim_{t \rightarrow \infty} \mathbb{P}\left(\frac{L(t) - 2t}{t^{1/3}} \leq x\right) = F(x) := \exp\left(-\int_x^\infty (s - x)q^2(s)ds\right) \quad (2)$$

where $q(x)$ is the (unique) solution[24] to the Painleé II equation

$$q'' = 2q^3 + xq \quad (3)$$

satisfying the condition

$$q(x) \sim -Ai(x), \quad x \rightarrow +\infty \quad (4)$$

*Department of Mathematics, University of Michigan, Ann Arbor, MI 48109 USA, E-mail: baik@umich.edu

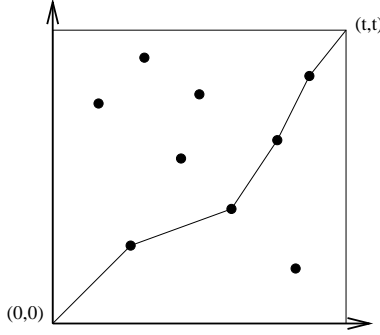


Figure 1: Poisson points and a longest up/right path from $(0,0)$ to (t,t) . In this example, $L(t) = 4$.

where Ai denotes the Airy function. Convergence of all the moments is also proved in the same paper. The limiting distribution $F(x)$ is called the Tracy-Widom distribution, which will be discussed further in section 3.

This article intends to provide some motivations and applications of the above Poisson last passage percolation model, and discuss the result (2).

2 Motivations and Applications

Random permutation

Given a realization of the Poisson points in the square $(0,t) \times (0,t)$, suppose one label them as $(x_i, y_{\pi(i)})$, $i = 1, \dots, N$ such that $x_1 < x_2 < \dots < x_N$. Note that with probability 1 no two points have the same x - or y -coordinate. Then the indices of the y -coordinates of the points generate a permutation π . Moreover, an up/right path is mapped to an increasing subsequence of the corresponding permutation. In the example of Fig. 1, the associated permutation is 629473518 and the increasing subsequence corresponding to the indicated up/right path is 2358. Therefore denoting the length of the longest increasing subsequence of S_N by L_N and recalling the property of the Poisson process discussed earlier, we find that

$$\mathbb{P}(L(t) = \ell) = \sum_{N=0}^{\infty} \frac{e^{-t^2} (t^2)^N}{N!} \mathbb{P}(L_N = \ell), \quad \ell \in \mathbb{Z}. \quad (5)$$

Using this formula and the so-called de-Poissonization lemma[25] one can extract from the result (2) that[4]

$$\lim_{N \rightarrow \infty} \mathbb{P}\left(\frac{L_N - 2\sqrt{N}}{N^{1/6}} \leq x\right) = F(x). \quad (6)$$

The problem of finding various statistics of the longest increasing subsequence of a random permutation in the large N limit has been known as Ulam's problem since early 1960's. The

Poisson version of the model as in the last passage percolation model above is sometimes called the Hammersley's process. See e.g. [1, 16] for more history of this combinatorial problem and its applications.

Plancherel measure

A partition of N is a sequence of integers $\lambda = \{\lambda_j\}_{j=1}^N$ such that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$ and such that the sum of λ_j 's is N . Given a partition λ of N , let d_λ denote the number of standard Young tableaux of shape λ (see e.g. [46] for definition). It is a basic result of representation theory of the symmetric group that d_λ is the dimension of the irreducible representation of S_N parameterized by λ . Hence the sum of d_λ^2 over all partitions λ of N is equal to N . In view of this identity, a natural probability on the set of partitions of N is the Plancherel measure defined by

$$\mathbb{P}(\lambda) = \frac{d_\lambda^2}{N!}. \quad (7)$$

Now the famous Robinson-Schensted[45] algorithm states that one can uniquely associate a pair of Young tableaux of a partition λ of N to each permutation $\pi \in S_N$. Moreover $L_N(\pi)$ is equal to the largest part λ_1 of the corresponding partition. This implies that the distribution of the largest part λ_1 of a random partition of N taken according to the Plancherel measure (7) is precisely equal to the distribution of L_N of a random permutation. Hence the result (6) yields the limiting law for λ_1 .

The asymptotic statistics of other parts $\lambda_2, \lambda_3, \dots$ have also been studied. See e.g. [32, 5, 37, 14, 27, 6, 49].

Polynuclear growth (PNG) model

Consider a one-dimensional flat substrate. Suppose that there occur random nucleation events, which is a Poisson process in the space-time plane. If a nucleation occurs at (x_0, t_0) , an island of height 1 with zero width is created at position x_0 at time t_0 . As time increases, the island grows laterally in both directions with speed 1 while keeping its height. Often two growing islands of same height collide. In that case they form one island and the edges of the amalgamated island keep growing with the same speed 1. Note that nucleations can occur on top of an existing island, and hence new islands may be created on an old island. Let $h(x, t)$ be the height of PNG model at position x at time t . Thus $h(x, t)$ is a piecewise constant function in x for fixed t . An example of the graph of $h(x, t)$ is in Fig. 2.

One could impose various initial conditions. For now, suppose that the nucleation events occur only for $|x| < t$. A different way of describing this condition is that at time $t = 0$, the substrate

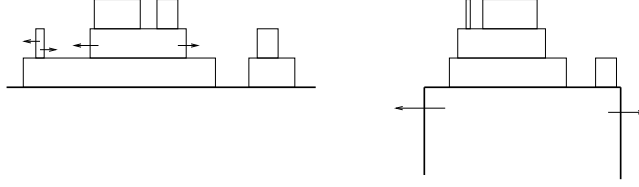


Figure 2: A snapshot of the height function of a PNG model. The picture on the right is of droplet initial condition.

satisfies $h(0,0) = 0$ and $h(x,0) = -\infty$ for all $x \neq 0$, and as time increases, the base substrate itself grows laterally of speed 1 and nucleations occur only on top of the base substrate (and of course on top of islands on the base substrate). This condition is called the *droplet case*. See the picture on the right in Fig. 2 for an example.

An observation by Prähofer and Spohn[40] is that the PNG model with the droplet initial condition can be mapped to the above Poisson last passage percolation model. Imagine the space-time plane in which the Poisson points corresponding to nucleation events are marked by dots. Due to the droplet condition, the dots are in the forward-light-cone $|x| < t$, and as the growth speed of islands is 1, the height at (x_1, t_1) depends only on the nucleation events in the backward-light-cone $|x - x_1| < t_1 - t$. Therefore, $h(x_1, t_1)$ depends only on the Poisson points in a rectangle of area $|t_1^2 - x_1^2|/2$. By rotating the coordinates by $-\pi/4$, and by re-scaling, one ends up with a Poisson process of rate 1 in a square of area $|t_1^2 - x_1^2|/2$. Moreover, one can observe that the height $h(x_1, t_1)$ is precisely equal to the length of the longest up/right path.

By re-interpreting the result (2) in terms of the PNG model using this identification, Prähofer and Spohn[40] found the following result: for fixed $|c| < 1$,

$$\lim_{t \rightarrow \infty} \mathbb{P} \left(\frac{h(ct, t) - \sqrt{2(1-c^2)}t}{((1-c^2)/2)^{1/6} t^{1/3}} \leq x \right) = F(x). \quad (8)$$

Note that the super-diffusive scaling $t^{1/3}$ is due to the fact that the height at a position strongly depends on the heights at the neighboring positions. Actually it has been believed that the exponent $1/3$ should be universal for a wide class of one-dimensional random growth models as long as the spatial correlation is not too weak (see e.g. [13]). Such models are said to be in the KPZ universality class. In the famous work[30], Kardar, Parisi and Zhang introduced a nonlinear stochastic differential equation for the surface height $h(x, t)$ as a continuum model for this class of random growth models, and renormalization group analysis has suggested that the scaling exponent should be $1/3$ for one-dimension case. The PNG model now plays the role of the unique growth model for which the $1/3$ exponent can be rigorously proved. Moreover, one can even establish the limiting distribution as in (8).

One might ask what happens if the initial condition is changed. It turned out that while the scaling exponent is the same, the limiting distribution may change. For instance, instead of the droplet condition, consider the flat initial condition: the initial substrate is \mathbb{R} and nucleation events could occur at any position on it. Then for any fixed position x , one finds (see [10, 40])

$$\lim_{t \rightarrow \infty} \mathbb{P} \left(\frac{h(x, t) - \sqrt{2}t}{2^{-1/6}t^{1/3}} \leq x \right) = \exp \left(-\frac{1}{2} \int_x^\infty q(s) ds \right) \cdot F(x)^{1/2}. \quad (9)$$

As a second example, consider a PNG model on a half line $\{x : x \geq 0\}$. In this case, one could imagine the situation such that extra nucleation events occur *at* $x = 0$, which corresponds to a 1-dimensional Poisson process in time at $x = 0$. In other words, there is excessive creation of islands at the origin. Then depending on the rate α of the creation of islands at the origin, the height function could have different properties. Indeed if α is big, then the height at $x = 0$ is dominated by the creation of islands at the origin, while if α is small, then it is likely that the islands created at the origin have little effect for the height at $x = 0$. Thus one expects a transition of $h(0, t)$ in α , which is actually proved in [10, 42]. On the other hand, at $x = ct$ for a fixed $c > 0$, the height is proved to have the fluctuation law given by $F(x)$ for the choices $\alpha = 0, 1$ [44]. The authors of [44] also computed height fluctuation for the transitional case for $\alpha = 0, 1$ when $x \sim t^{2/3}$ in terms of a Fredholm determinant. It is yet to be seen to obtain a Painlevé II type formula for this determinant. For further references in this direction, see e.g. [3, 10, 11, 12, 41, 42, 43]. We note that all these different initial conditions have combinatorial meanings on random permutations and Plancherel measure.

Particle/anti-particle process

A different description of the PNG model[43] is to regard the right edge of an island as a particle and the left edge of an island as an anti-particle. Hence there are right-moving particles and left-moving anti-particles on the real line. Creation of island corresponds to creation of particle and anti-particle pair, and the fact that two islands stick together when they meet implies that upon colliding, particle and anti-particle annihilate each other. In this dynamic picture, the height function is now equal to the total number of the particles and the anti-particles that have crossed the given position up to the given time.

Random vicious walks

The combinatorics of the longest increasing subsequence and the Plancherel measure have an interpretation as non-intersecting paths, which is sometimes called vicious walks[20]. See e.g. [21, 22, 23, 29] for reference.

3 Random matrix and universality

The Tracy-Widom distribution $F(x)$ in (2) also appears in a totally different subject; random matrix theory. The main interest in the random matrix theory is the limiting statistics of the eigenvalues as the size of matrix tends to infinity. Random matrix theory has very diverse applications in both mathematics and physics from the spectrum of heavy nuclei to the zeros of Riemann-zeta function (see e.g. [35, 17, 31]).

Of special interest is the Gaussian unitary ensemble (GUE) which is the set of $N \times N$ Hermitian matrices H equipped with the probability measure[35]

$$\frac{1}{Z_N} e^{-N \text{tr}(H^2)} dH \quad (10)$$

where dH is the Lebesgue measure and Z_N is the normalization constant. In 1994 Tracy and Widom[48] considered the limiting distribution of the largest eigenvalue $\xi_{\max}(N)$ of $N \times N$ Hermitian matrix taken from GUE and found that

$$\lim_{N \rightarrow \infty} \mathbb{P}((\xi_{\max}(N) - \sqrt{2})\sqrt{2}N^{2/3} \leq x) = F(x) \quad (11)$$

where $F(x)$ is precisely the same function in (2). In other words, upon proper scaling, the largest eigenvalue of a random Hermitian matrix taken from GUE and the length of the longest up/right path in the Poisson last passage percolation model have the same limiting law.

It should be mentioned that the analysis of obtaining (2) and (11) are independent. Especially it is not found whether there is a direct relation between $L(t)$ and $\xi_{\max}(N)$ for finite t and N . Only in the limit $t \rightarrow \infty$ and $N \rightarrow \infty$, two seemingly different quantities have the same limit after independent computations. A framework to understand this might be central limit theorem. In the classical central limit theorem, the sum of n independent identically distributed random variables converges, after proper scaling, to the Gaussian distribution as $n \rightarrow \infty$, irrelevant to the detail of the random variable. The results (2) and (11) have the similar feature that the two different ‘random variables’ share the same limiting distribution. So one may expect that the Poisson percolation model and GUE are two instances of models to which a nonlinear version of central limit theorem is applied. It is however not known whether there is indeed such a natural nonlinear version of the central limit theorem with proper general setting that includes the Poisson percolation model and the GUE. But there are some ‘patches’ of universality results in random matrix theory, and in the remaining of this section, we discuss some of them.

Universality of random matrices

A generalization of GUE is the set of Hermitian matrices with the density

$$\frac{1}{Z_N} e^{-N \text{tr} V(H)} dH \quad (12)$$

where V is an analytic function with sufficient decay properties. We remark that then the eigenvalues $\xi_1 \geq \xi_2 \geq \dots \geq \xi_N$ have the density

$$\frac{1}{Z'_n} \prod_{1 \leq i < j \leq N} |\xi_i - \xi_j|^2 \cdot \prod_{j=1}^N w(x_j) \quad (13)$$

for some constant Z'_N and $w(x) = e^{-NV(x)}$. This can be thought of as a Coulomb gas of particles in \mathbb{R} with logarithmic repulsion and external potential w . For general V , the local statistics of the eigenvalues in the middle of the limiting density of states are found to be independent of V (see [39, 15, 18]). Moreover, by using the results of [18] and [34], one can also deduce that the limiting distribution of the largest eigenvalue of random Hermitian matrix taken according to the probability (12) is generically given by the Tracy-Widom distribution $F(x)$. The special case of $V(x) = x^4 + tx^2$ was obtained by Bleher and Its[15].

The GUE has an alternative definition and a different generalization. Namely, GUE is the set of Hermitian matrices with independent, except for the Hermitian condition, complex Gaussian entries. It is direct to check from this definition that the density of the matrix is precisely (10). A natural generalization is then the random Hermitian matrices of independent entries which are not necessarily Gaussian. This is called the Wigner matrix[35]. For Wigner matrix, the largest eigenvalue is found to have the same limit $F(x)$ again[46] (see also [28] for a result regarding the eigenvalues in the middle of the limiting density of states). Hence the limiting law (11) still holds true for two different generalizations of GUE, one of density function (12) and the other of independent entries.

Growth models and discrete orthogonal polynomial ensembles

In addition to the Poisson percolation model, there are a few more isolated examples of percolation models for which the scaling limit (2) can be obtained. Consider the lattice sites \mathbb{N}^2 . Suppose that we assign a random variable $X(i, j)$ at each site $(i, j) \in \mathbb{N}^2$. We further assume that $X(i, j)$ are independent and identically distributed. Consider an up/right path π from the site $(1, 1)$ to (M, N) , which is a collection of neighboring sites $\{(i_k, j_k)\}$ such that $(i_{k+1}, j_{k+1}) - (i_k, j_k)$ is either $(1, 0)$ or $(0, 1)$. Let $\Pi(M, N)$ denote the set of up/right paths from $(1, 1)$ to (M, N) , and define

$$L(M, N) := \max_{\pi \in \Pi(M, N)} \left\{ \sum_{(i, j) \in \pi} X(i, j) \right\}. \quad (14)$$

If $X(i, j)$ are positive random variables, an interpretation is that $X(i, j)$ is the passage time at the site (i, j) , and $L(M, N)$ is the last passage time to travel from $(1, 1)$ to (M, N) along an admissible up/right path. The Poisson percolation model is a continuum version of this more general directed last passage percolation model.

For general random variables $X(i, j)$, the scaling limit law (2) is an open problem, but when the random variable $X(i, j)$ is either geometric or exponential, (2) is proved[26]. Also if the definition of the up/right paths is modified, there are a few more isolated cases for which (2) is obtained (see e.g. [48, 27, 2]). However all the cases such that (2) is proved share the common feature that they all have an interpretation as a version of the longest increasing subsequence of a random (generalized) permutation, and all of them have the same algebraic structure (see e.g. [38, 9]). Hence it is an open question to prove the universality result (2) for a general random variable $X(i, j)$ which does not have such a structure.

The geometric percolation model above has an alternative representation. Consider the density function on the set of particles $\xi_1 > \xi_2 > \cdots > \xi_N \geq 0$, $\xi_j \in \mathbb{N} \cup \{0\}$, given by (13). The only change is that the ‘particles’ ξ_j lie on a discrete set instead of a continuous set. Sometimes this is called the ‘discrete orthogonal polynomial ensemble’, while (13) with continuous weight w is called the ‘continuous orthogonal polynomial ensemble’. A result of Johansson[26] is that for the geometric percolation model, $L(M, N)$ has the same distribution (except for a minor translation change) as the ‘largest particle’ ξ_1 in the discrete orthogonal polynomial ensemble with the special choice $w(x) = \binom{x+M-N}{x} q^x$ (assuming that $M \geq N$).

Hence one may wonder whether the discrete orthogonal polynomial ensemble with general weight w have universal properties just like its continuous weight counterpart (13). This is indeed proved to be the case in [7, 8] for a wide class of discrete weight w . The analysis extends the Deift-zhou steepest-descent method[19] of Riemann-Hilbert analysis to the discrete interpolation setting (see also [33, 36]). For general discrete weights w , it is not clear if discrete orthogonal polynomial ensembles have any percolation-type interpretation, but the universality of (13) for both continuous and discrete weights provides a linkage of the result (11) for GUE and the result (2) for the geometric percolation model.

Acknowledgments

The author would like to thank Percy Deift and Michio Jimbo for kindly inviting him to the special session of integrable systems in the International Congress on Mathematical Physics. The work is supported in part by NSF Grant # DMS-0208577.

References

- [1] D. Aldous and P. Diaconis, *Bull. Amer. Math. Soc. (N.S.)* **36**(4), 413–432 (1999).
- [2] J. Baik, *Comm. Pure Appl. Math* **53**(11), 1385–1410 (2000).
- [3] J. Baik, in **Contemporary Mathematics** **326**, 1–21, AMS, 2003.
- [4] J. Baik, P. A. Deift and K. Johansson, *J. Amer. Math. Soc.* **12**(4), 1119–1178 (1999).
- [5] J. Baik, P. A. Deift and K. Johansson, *Geom. Funct. Anal.* **10**(4), 702–731 (2000).
- [6] J. Baik, P. A. Deift and E. Rains, *Comm. Math. Phys.* **223**(3), 627–672 (2001).
- [7] J. Baik, T. Kriecherbauer, K. T.-R. McLaughlin and P. Miller, *Int. Math. Res. Not.* **15**, 821–858 (2003).
- [8] J. Baik, T. Kriecherbauer, K. T.-R. McLaughlin and P. Miller, in preparation.
- [9] J. Baik and E. Rains, *Duke Math. J.* **109**(1), 1–65 (2001).
- [10] J. Baik and E. Rains, *Duke Math. J.* **109**(2), 205–281 (2001).
- [11] J. Baik and E. Rains, in **Math. Sci. Res. Inst. Publ.**, **40** 1–19, Cambridge Univ. Press, Cambridge, 2001.
- [12] J. Baik and E. Rains, *J. Stat. Phys.* **100**(3–4), 523–541 (2000).
- [13] A. Barabási and H. E. Stanley, **Fractal concepts in surface growth**, Cambridge University Press, Cambridge 1995.
- [14] A. Borodin, A. Okounkov and G. Olshanski, *J. Amer. Math. Soc.* **13**(3), 481–515 (1999).
- [15] P. Bleher and A. Its, *Ann. of Math. (2)* **150**(1), 185–266 (1999)
- [16] P. A. Deift, *Notices Amer. Math. Soc.* **47**(6), 631–640(2000).
- [17] P. A. Deift, **Orthogonal polynomials and random matrices: a Riemann-Hilbert approach**, AMS, RI 1999.
- [18] P. A. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides and X. Zhou, *Comm. Pure Appl. Math.* **52**(11), 1335–1425 (1999).
- [19] P. A. Deift and X. Zhou, *Ann. Math. (2)* **137**, 295–368 (1993).

- [20] M. Fisher, *J. Statist. Phys.* **34**(5-6), 667–729 (1984).
- [21] P. Forrester, *J. Phys. A* **23** 1259–1273 (1990).
- [22] P. Forrester, *J. Phys. A* **34**(31), L417–L423 (2001).
- [23] A. Guttmann, A. Owczarek and X. Viennot, *J. Stat. Phys.* **102**(5-6), 1085–1132 (2001).
- [24] S. Hastings and J. McLeod, *Arch. Rational Mech. Anal.* **73**, 31–51 (1980).
- [25] K. Johansson, *Math. Res. Lett.* **5**(1-2), 63–82 (1998).
- [26] K. Johansson, *Comm. Math. Phys.* **209**(2), 437–476 (2000).
- [27] K. Johansson, *Ann. of Math.* **153**, 259–296 (2001).
- [28] K. Johansson, *Comm. Math. Phys.* **215**(3), 683–705(2001).
- [29] K. Johansson, *Probab. Theory and Related Fields* **123**, 225–280 (2002).
- [30] M. Kardar, G. Parisi and Y. Z. Zhang, *Phys. Rev. Lett.* **56**, 889–892 (1986).
- [31] N. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, AMS, RI 1999.
- [32] S. Kerov, *Funct. Anal. Appl.* **27**, 104–117 (1993).
- [33] T. Kriecherbauer, K. T.-R. McLaughlin and P. Miller, to appear in *Ann. of Math. Stud.*
- [34] A. B. J. Kuijlaars and K. T.-R. McLaughlin, *Comm. Pure Appl. Math.* **53**(6), 736–785 (2000).
- [35] M. Mehta, **Random matrices**, Academic press, San Diago 1991.
- [36] P. Miller, *Int. Math. Res. Not.* **8**, 383–454 (2002).
- [37] A. Okounkov, *Int. Math. Res. Not.* **20**, 1043–1095 (2000).
- [38] A. Okounkov, *Selecta Math. (N.S.)* **7**(1), 57–81 (2001).
- [39] L. Pastur and M. Shcherbina, *J. Stat. Phys.* **86**(1-2), 109–147 (1997).
- [40] M. Prähofer and H. Spohn, *Physica A* **279**, 342–352 (2000).
- [41] M. Prähofer and H. Spohn, *Phys. Rev. Lett.* **84**, 4882 (2000).
- [42] M. Prähofer and H. Spohn, in **Progr. Probab.** **51**, 185–204, Birkhäuser, Boston, MA, 2002.

- [43] M. Prähofer and H. Spohn, <http://xxx.lanl.gov/abs/cond-mat/0212519>, preprint (2002).
- [44] T. Sasamoto and T. Imamura, <http://xxx.lanl.gov/abs/cond-mat/0307011>, preprint (2003).
- [45] C. Schensted, *Canad. J. Math.* **13**, 179–191 (1961).
- [46] A. Soshnikov, *Comm. Math. Phys.* **207**(3), 697–733(1999).
- [47] R. P. Stanley, **Enumerative Combinatorics**, Cambridge University Press, Cambridge, United Kingdom 1999.
- [48] C. Tracy and H. Widom, *Comm. Math. Phys.* **159**, 151–174 (1994).
- [49] C. Tracy and H. Widom, *Probab. Theory Related Fields* **119**(3), 350–380 (2001).
- [50] H. Widom, *Int. Math. Res. Not.* **9**, 455–464 (2002).

